Why does Kruskal's algorithm work?

We show that the spanning tree returned by Kruskal's algorithm has minimum cost for the special case where no two edges have the same weight.